கணினி அகராதியியல்

முனைவர் ச. இராஜேந்திரன்
மொழியியல் துறை
தமிழ்பல்கலைக்கழகம்
தஞ்சாவூர்

சொல்லியியல்

அகராதி ஆய்வில் கவனம் செலுத்துகிற மொழியியலின் ஒரு பகுதிதான் சொல்லியல் (Lexicology). இது அகராதியின் இயலுபு, அமைப்பு மற்றும் அதன் அலகுகளின் குணங்களை உருப்படுத்தம் செய்யும்; இயற்கை மொழிகளின் மாதிரிகளில் அகராதியின் பங்கை விளக்கும். கணினி சொல்லியல் சொல் ஆய்வில் கணினியின் பயன்பாட்டின் நோக்கை உள்ளடக்குகின்றது. குறிப்பாகப் பின்வருவனவற்றில் கவனக்குவிப்பு செய்கின்றது:

அகராதியின் கணினி உருப்படுத்தத்தில்
சொல் தரவில் செய்யப்படும் கணினிச் செயல்பாடுகள்
ஒரு நோக்கில் கணினி அகராதிகளுக்கிடையே உள்ள உறவுகள், மற்றொரு நோக்கில்
பல இயற்கை மொழி ஆய்வு (NLP) ஒழுங்குமுறையின் கூறுகளுக்கிடையே உள்ள உறவுகள்

அகராதியியல்

அகராதியியல் சொற்களைச் சேகரிப்பது மற்றும் அவை எவ்வாறு பயன்படுத்தப்படுகின்றது என்பதில் கவனம் செலுத்துவதாகப் பொதுவாகக் கருதப்படுகின்றது. அகராதியியல் உற்பத்திப்பொருள்களில் பல அச்சடிக்கப்பட்ட அகராதிகள் மட்டுமன்றி அகரவரிசை அட்டவணை, சொல்லடைவு, தொடரடைவு, கலைச்சொற்கள் போன்றவையும் அடங்கும்.

அகராதியியல் என்றால் என்ன?

குறிப்பிடத்தகுந்த சொற்களையும் கலைச்சொற்களையும் தொடர்பான செய்திகளுடன் கண்டுபிடிப்பது, அம்மாதிரியான செய்திகளைத் தொகுப்பது, குறியாக்கம் செய்வது, முன்னிலைப்படுத்தித் தருவது, அம்மாதிரியான செய்திகளைப் பரவச்செய்தல் மற்றும் வழங்குதல், அந்தச் செய்திகளின் பயன்பாட்டை ஆய்தல் மற்றும் அதன் தரத்தை அளத்தல்

அகராதி உருவாக்கல்

அகராதி உருவாக்கல் என்பது பன்நோக்கு செயற்பாங்காகும்: இது குறைந்தது பின்வரும் செயற்பாடுகளை உள்ளடக்கும்:

- மொழியைப்/மொழிகளைப் புரிந்து கொள்ளுதல் (Understanding the language(s); அமைப்பு, செயற்பாடு, பொருண்மை மற்றும் சமுதாய-பண்பாட்டு அடிப்படையில் மொழியின் அமைப்பைப் புரிந்து கொள்ளுதல்.
- செய்திகளை அமைப்பபாக்கம் செய்தல் (Structuring the information): அகராதிச் சொல்லின் கீழ் மேற்சொன்ன செய்திகளை ஒரு அமைப்பொழுங்கில் தருதல்.
- செய்திகளை உள்ளீடு செய்தல் (தரவு மையத்தைத் தொகுத்தல்) (Inputting the information (compiling the lexical database): இயல்பாக ஒரு ஆண்டிற்குச் செய்திகளை உள்ளீடு செய்தல், இது ஆரம்பத்திலேயே செய்யப்பட வேண்டும், தொடர்ந்து செய்யப்பட வேண்டும்.
- பரிசோதித்து மேன்மைப்படுத்துதல் (Checking and refining): சொல் தரவு மையத்திலுள்ள செய்திகளைப் பரிசோதனை செய்து மேன்மைப் படுத்தல்.
- தரவைப் பயன்படுத்தல் (Manipulation of the data): தரவைச் சொற்களங்களைப் பெறுவதற்காகவும், தலைகீழான செயல்பாடுகளைச் செய்வதற்காகவும் பயன்படுத்தல்.
- வெளியீடு: வடிவமைப்பைத் தீர்மானித்து வேண்டிய மாற்றங்களைச் செய்தல்.

- அச்சிடல்.
- வணிகம் செய்தல் மற்றும் வினியோகித்தல்.

அகராதி வகைகள்

- ஒரு மொழி, இருமொழி, பன்மொழி அகராதிகள்
- பொருட்புல அகராதிகள் மற்றும் கருத்துரு அடிப்படையிலான அகராதிகள்
- தனிப்பட்ட வகை (கொச்சை மொழி, சொல்மூலம் அல்லது சொல் வரலாறு)
- கலைக்களஞ்சியம் (உண்மை உலகத்தின் குறிப்புரை)
- மொழிபெயர்ப்பு (மொழிபெயர்ப்பு நிகரிகள்)
- பேருந்தரவுகள் (தொடரடைவு)
- சொற்றொகை (சொற்களம் அடிப்படையிலான சொற் பட்டியல்)
- பொருள்விளக்கச் சொற்கோவை (பழைய, அரிதான சொற்கள்)
- பொருள்களுக்குப் பெயரிடுதல்
- சொல்லடைவு
- சொல், கலைச்சொல் நிகழ்வெண்

அகராதியியலும் மொழியியலும்

மொழியியல் அகராதியியலாரின் அகராதியைச் செம்மைப்படுத்தப் பின்வரும் வழிகளின் பயன்படுகிறது:

- பொருண்மை வேறுபாடுகளின் ஒழுங்கமைப்பைக் கண்டுகொள்ள
- தொடரியல் பண்புகளையும் பொருண்மையியல் பண்புகளையும் பொருத்த
- ஒரே கருத்துரு வகுப்பைச் சார்ந்த சொல் பொருண்மைகளை விவரணை செய்யப் பொதுவான பொருண்மைச் சட்டத்தைக் கையாள

கணினி அகராதியியல்

கணினி மொழியியலாருக்கும் அகராதியியலாருக்கும் இடையே உள்ள பரிவர்த்தனை அகரவரிசை அட்டவணை தயாரிக்கவும், பெரிய தரவுகளிலிருந்து தொடரடைவு தயாரிக்கவும் வேண்டி எளிய பயன்பாட்டுக்கருவிகள் தயாரிக்கும் எல்லைக்குட்பட்டதாய் இருந்தது.

கணினியைப் பயன்படுத்திச் செய்திகளைக் பயன்படுத்துபவரைக் கவரத்தக்க விதத்தில் திட்டமிட்டுதல், குறியாக்கம் செய்தல், அகராதிச் செய்திகளைப் பரப்புதல்
அகராதித் தரவு மூலங்களை நடைமுறைப்படுத்தல், ஊக்குதல், மற்றும் ஒன்றுசேர்த்தல்
கணினியால் படிக்கப்படும் அகராதிகளை உருவாக்கக் கருவிகளைத் தருதல், ஒப்பிடுதல் மற்றும் திறனாய்தல்
பெரிய தரவுகளிலிருந்து உண்மையான மொழிப் பயன்பாடு அடிப்படையில் சான்றுகளைப் பிரித்தெடுக்கும் அகராதி வழிமுறைகளின் பயனை ஆய்தல்
வருணனைத் திட்டங்களைப் பரிசோதிக்கப் பயன்படும் தரவுகளைத் தேர்ந்தெடுத்தல் மற்றும் சேகரித்தல்
ஆய்வுகளின் மேம்பாடு மற்றும் மாற்றத்துடன் திறன் ஆய்வு செய்தல்

கணினி அகராதியியல் பின்வருவனவற்றை உள்ளடக்கும்

அகராதி உருவாக்க வேலைகளுக்கு உதவிபுரிய வழிமுறைகள் மற்றும் கருவிகளைத் தயாரித்தல்
பல மூலங்களிலிருந்து சான்றுகளைத் தயாரித்தல்
பொருத்தமான மொழியியல் செய்திகளின் தரவு அடிப்படையைப் பதிவுசெய்தல்
அகராதிப் பதிவுகளைச் சீரமைத்தல்
அகராதி உற்பத்திப்பொருள்களைப் பரவச்செய்தல்

கணினி மொழியியலார் அகராதியியலாரின் திறமைகளை உணர்ந்திருக்க வேண்டும். வெளியீட்டாளர்கள் வைத்திருக்கிற தரவுகள் இயற்கை மொழி

ஆய்வுக்காகப் பெரிய அகராதிகளை உருவாக்கச் செய்கின்ற பங்களிப்பைப் பெரிதுபடுத்தவேண்டும். கணினி சொலிலியலும் அகராதியியலும் பிற பாடங்களுடனும் தொடர்புடையது, குறிப்பாகக் கணினி அறிவியல், கோட்பாட்டு மற்றும் கணினி மொழியியல் மற்றும் அகராதியியலுடன் தொடர்புடையது. மொழியியலில் கணினி தொழில்நுட்பத்தை அறிமுகப்படுத்தியது கணினி மொழியியலில் புதிய ஆய்வுக்களங்களைத் தோற்றுவித்தது.

பனுவல் பெருந்தரவை ஆய்தல்

எந்தவிதமான செய்திகளைக் கணினி படிக்கும் பனுவல்களிலிருந்து பெறலாம்? பனுவல் என்றால் என்ன?

மொழியியல் அடிப்படையில் பனுவல் என்பதைப் பொருண்மையில், தொடரியல், பயன்வழியல், அடிப்படையில் ஒன்றோ அதற்கு மேலோ உள்ள சொற்குழுமங்களின் சீரான முழுமை எனலாம்.

இரண்டாவது கேள்வி, இவ்வாறான பனுவல் இருந்தால் அகராதியியல் அடிப்படையில் எவ்வகையிலான பயனுள்ள வெளியீடுகளைப் பனுவல் ஆய்விலிருந்து எதிர்பார்க்கலாம்? பின்வருமாறு உருப்படுத்தம் செய்யலாம்:

விளக்கச் செய்தி (Definitional Information) அல்லது அகராதிச் சொற்களின் அகப்பொருண்மை அமைப்பு பற்றிய செய்தி
குறிப்பீட்டுச் செய்தி (Referential Information) அல்லது மொழியியலுக்கு அப்பாற்பட்ட கருத்துருக்கள் மற்றும்/அல்லது பொருட்களின் அகராதியாக்கம் பற்றிய செய்தி
கருத்துருச் செய்தி (Conceptual Infomaration) அல்லது சொற்களின் பொருண்மைகளுக்கு இடையே உள்ள உறவுகள் பற்றிய செய்தி சேர்ந்துவருகைச்

செய்தி (Collocational Information) அல்லது இணைந்து வரும் சொற்களின் பொருள் பற்றிய செய்தி.

* * * * *